

# Protecting coherence in Optimal Control Theory: state dependent constraint approach

José P. Palao,<sup>1</sup> Ronnie Kosloff,<sup>2</sup> and Christiane P. Koch<sup>3</sup>

<sup>1</sup>*Departamento de Física Fundamental II, Universidad de La Laguna, La Laguna 38204, Spain*

<sup>2</sup>*Institute of Chemistry and The Fritz Haber Research Center, The Hebrew University, Jerusalem 91904, Israel*

<sup>3</sup>*Institut für Theoretische Physik, Freie Universität Berlin, Arnimallee 14, 14195 Berlin, Germany*

(Dated: March 6, 2008)

Optimal control theory is developed for the task of obtaining a primary objective in a subspace of the Hilbert space while avoiding other subspaces of the Hilbert space. The primary objective can be a state-to-state transition or a unitary transformation. A new optimization functional is introduced which leads to monotonic convergence of the algorithm. This approach becomes necessary for molecular systems subject to processes implying loss of coherence such as predissociation or ionization. In these subspaces controllability is hampered or even completely lost. Avoiding the lossy channels is achieved via a functional constraint which depends on the state of the system at each instant in time. We outline the resulting new algorithm, discuss its convergence properties and demonstrate its functionality for the example of a state-to-state transition and of a unitary transformation for a model of cold  $\text{Rb}_2$ .

PACS numbers:

## I. INTRODUCTION

Coherent control utilizes the wave properties of matter to steer a quantum dynamical process to a desired outcome. The source of control is interference, constructive to achieve the goal and destructive to eliminate unwanted consequences [1, 2]. The agents of control are external fields, in particular electromagnetic fields. The experimental and theoretical challenge lies in identifying these control fields. The present study is aimed at finding control fields which are constrained to limit the damage which the control field may cause to the controlled system.

Theoretically, the control problem can be formulated as an inversion: finding the field subject to the quantum dynamics which leads to the desired outcome. Optimal control theory (OCT) has been developed as a tool to address this problem [3, 4]. It can be formulated starting from a variational ansatz [3] or using Krotov's method [5, 6]. Recently the Krotov's method has been extended to include a strict limitation on the spectrum of the optimized field [7].

The most well studied task in OCT has been the goal of a state-to-state transition. Given an initial state  $\psi_{ini}$  and a closed quantum system, the field needs to be found which drives the system to a specific final state  $\psi_{fin}$ . This task has been shown to be completely controllable [8, 9] if the fields are not restricted. Moreover the control landscape is favorable composed of flat ridges such that the climb in the gradient direction will lead to one of the many possible solutions [10].

A more involved control task is to optimize the expectation value of an operator at a final time,  $\langle \hat{\mathbf{B}}(t_{fin}) \rangle$ . This task can be formulated in the framework of open quantum systems. The OCT approach yields an iterative solution to the inversion problem which is based on propagating the system density operator  $\hat{\rho}_S(t)$  forward

in time and the target operator  $\hat{\mathbf{B}}(t)$  backward in time [11, 12, 13].

The prospect of quantum computing has posed an even more complex control problem: imposing a unitary transformation  $\hat{\mathbf{U}}$  on a subset of quantum states which act as the quantum register. The unitary transformation carries out a specific computational task. This control task is equivalent to  $N$  simultaneous state-to-state transformations [6, 14, 15]. The solution of the iterative set of equations has been shown to become exponentially more difficult with the size  $N$  of the unitary transformation [6]. These findings are in accordance with a very complex control landscape [16].

A further step up in complexity is the task of imposing a unitary transformation under dissipative conditions. This task emerges in the quantum governor [17], in quantum information processing and it is a traditional task in nuclear magnetic resonance (NMR) spectroscopy [18].

In any practical implementation the positive task of obtaining the final goal has to be weighted by possible negative consequences. For example control fields of high intensity can damage the system by causing ionization or dissociation. A remedy for this problem consists in restricting the population in certain lossy excited state manifolds. This task has been the motivation for the development of local control theory (LCT) [19, 20]. LCT has been applied to lock unwanted electronic excitations [21] and recently to the problem of quantum information processing where avoiding population loss becomes crucial [22, 23].

However, OCT is more powerful than LCT and it is therefore desirable to incorporate constraints describing negative consequences of the control process into the algorithm. Such constraints depend on the state of the system at intermediate times [24, 25, 26, 27]. For example, the system can simply be restricted to remain in an "allowed" or to avoid a "forbidden" subspace during its evo-

lution. More elaborate examples include imposing a pre-defined path between an initial and a final state [26], and maximizing the expectation value of a time-dependent operator throughout the optimization time interval [25].

Previous OCT studies which impose state-dependent constraints [24, 25, 26] were performed for state-to-state optimization and are based on the variational approach. In the present work, an optimization algorithm including state-dependent constraints is obtained using the Krotov method for the state-to-state case as well as for unitary transformations. The Krotov method offers the advantage that the monotonic convergence of the algorithm can be ensured by the choice of the imposed constraints [6, 28]. A brief comparison with the algorithms using the variational ansatz [24, 25, 26] will be given for the state-to-state optimization.

The paper is organized as follows: The state-dependent constraints are formulated in Sec. II, and the resulting algorithm is presented for optimization of state-to-state transition and of a unitary transformation. A review of the Krotov method together with an outline of the derivation of the equations presented in Sec. II is given in the Appendix. Sec. III introduces a model example and illustrates optimization under state-dependent constraints for a state-to-state transition and for a unitary transformation. Our findings are compared to related approaches in Sec. IV. Finally, Sec. V concludes.

## II. FORMULATION OF STATE-DEPENDENT CONSTRAINTS IN THE KROTOV METHOD

### A. Optimization of a state-to-state transition

The dynamics of the system is governed by the time dependent Schrödinger equation

$$\frac{d}{dt}|\varphi(t)\rangle = -\frac{i}{\hbar}\hat{\mathbf{H}}[\epsilon(t)]|\varphi(t)\rangle, \quad (1)$$

where  $|\varphi(t)\rangle$  represents the state of the system at time  $t$ , and

$$\hat{\mathbf{H}}[\epsilon] = \hat{\mathbf{H}}_0 - \hat{\boldsymbol{\mu}}\epsilon(t), \quad (2)$$

is the system+control Hamiltonian.  $\hat{\mathbf{H}}_0$  denotes the field-free Hamiltonian,  $\epsilon(t)$  the semiclassical control field and  $\hat{\boldsymbol{\mu}}$  is a system operator describing the coupling between system and field.

The objective of the optimization is to find a field which drives the system from an initial state at  $t = 0$ ,

$$|\varphi(t=0)\rangle = |\varphi_0\rangle, \quad (3)$$

to a target subspace at time  $T$  representing the final time objective, such that a minimum (or maximum) expectation value of the time-dependent operator  $\hat{\mathbf{P}}(t)$  is maintained throughout the complete time interval  $[0, T]$ . The target subspace at time  $T$  is described by the projector  $\hat{\mathbf{D}}$ , e.g.,  $\hat{\mathbf{D}} = |\varphi_f\rangle\langle\varphi_f|$  for a single target state.

In OCT, these requirements are formulated as a functional which depends on the system state and the control, in such a way that an optimal field corresponds to an extremum of the functional. That functional can be expressed as a sum over terms related to the different conditions imposed on the system evolution.

### 1. The functional

The complete functional is obtained as a sum over functionals corresponding to the final time objective, to the state-dependent intermediate-time objective (or constraint), and to the constraint over the field.

The term corresponding to the objective at the final time  $T$ , the actual target, can be expressed as

$$J_0[\varphi_T, \varphi_T^*] = \lambda_0 \langle \varphi(T) | \hat{\mathbf{D}} | \varphi(T) \rangle, \quad (4)$$

where  $\lambda_0$  is a real parameter, which can be negative or positive, depending on whether the functional is minimized or maximized during the optimization.  $[\varphi_T, \varphi_T^\dagger]$  emphasizes the bilinearity of the functional with respect to the system state at time  $T$ . Other possibilities for expressing this term exist [6, 24], but the resulting optimization algorithms are very similar.

The state-dependent intermediate-time objective or constraint is represented by the functional,

$$J_b[\varphi, \varphi^\dagger] = \int_0^T g_b[\varphi, \varphi^\dagger] dt, \quad (5)$$

where  $g_b$  is taken to be

$$g_b[\varphi, \varphi^\dagger] = \lambda_b \langle \varphi(t) | \hat{\mathbf{P}}(t) | \varphi(t) \rangle. \quad (6)$$

$\lambda_b$  is a real parameter which can be positive or negative, as discussed later. More complicated dependences of  $g_b$  on the operator  $\hat{\mathbf{P}}(t)$  and on the state  $\varphi$  are conceivable.

To obtain a closed algorithm, the complete functional has to include a term depending on the field [5, 6],

$$J_a[\epsilon] = \int_0^T g_a[\epsilon] dt. \quad (7)$$

Generally,  $g_a$  can be written as

$$g_a[\epsilon] = \lambda_a(t) [\epsilon(t) - \epsilon_r(t)]^2, \quad (8)$$

where  $\epsilon_r$  denotes a reference field and  $\epsilon_r = 0$  corresponds to the common choice of minimizing the field energy.  $J_a[\epsilon]$  represents an intermediate-time objective, but one which does not depend on the state of the system.

The complete functional is given by

$$J[\varphi, \varphi^\dagger, \epsilon] = J_0[\varphi_T, \varphi_T^\dagger] + J_a[\epsilon] + J_b[\varphi, \varphi^\dagger]. \quad (9)$$

For simplicity, we omit the dependence of  $\varphi$  and  $\epsilon$  on time, except for the final time  $T$ . The optimization problem is now equivalent to the minimization or maximization of this functional. For that purpose, the Krotov

method is employed. Since the Krotov method operates with real functions, a complete presentation of the equations for this problem is somewhat cumbersome [6]. An outline of the derivation is given in the Appendix and only the final result is presented below.

## 2. The optimization algorithm

A guess field is denoted by  $\epsilon^{(0)}(t)$  and the corresponding state  $|\varphi^{(0)}(t)\rangle$  is given by the evolution Eq. (1) with the initial condition  $|\varphi(t=0)\rangle = |\varphi_0\rangle$ . In the Krotov method a new field  $\epsilon^{(1)}(t)$  which decreases (or increases) the functional value is obtained by the following equations: A new “state”  $|\chi\rangle$  is determined using the inhomogeneous equation

$$\frac{d}{dt}|\chi(t)\rangle = -\frac{i}{\hbar}\hat{\mathbf{H}}[\epsilon^{(0)}(t)]|\chi(t)\rangle + \lambda_b\hat{\mathbf{P}}(t)|\varphi^{(0)}(t)\rangle, \quad (10)$$

with the “initial” condition

$$|\chi(T)\rangle = -\lambda_0\hat{\mathbf{D}}|\varphi^{(0)}(T)\rangle, \quad (11)$$

cf. Eqs. (A14,A17) of the Appendix. It corresponds to the common OCT result modified by the inhomogeneous term  $\lambda_b\hat{\mathbf{P}}(t)|\varphi^{(0)}(t)\rangle$  which arises from the state-dependent constraint. The state  $|\chi\rangle$  is used to determine the new control field,

$$\epsilon^{(1)}(t) = \epsilon^{(0)}(t) - \frac{1}{\hbar\lambda_a(t)}\text{Im}\left\{\langle\chi(t)|\hat{\boldsymbol{\mu}}|\varphi^{(1)}(t)\rangle\right\}. \quad (12)$$

cf. Eq. (A22) of the Appendix, where  $\epsilon_r \equiv \epsilon^{(0)}$  was chosen. This is an implicit equation since the state  $|\varphi^{(1)}\rangle$  which depends on  $\epsilon^{(1)}$  appears in the right-hand side of Eq. (12). The numerical discretization of this implicit equation has been widely discussed for the homogeneous case (see for example Ref. [6]). The inhomogeneous term in Eq. (10) requires a modification of the time propagation method. A symmetrical propagation scheme is employed based on the diagonalization of the Hamiltonian in the interleaved time grid points,  $t_i + \Delta t/2$ . The inhomogeneous term is evaluated as

$$\lambda_b\left(\frac{\hat{\mathbf{P}}(t_{i+1})|\varphi^{(0)}(t_{i+1})\rangle + \hat{\mathbf{P}}(t_i)|\varphi^{(0)}(t_i)\rangle}{2}\right). \quad (13)$$

The iterative algorithm is constructed with  $\epsilon^{(1)}$  as input to the next step of the iteration and the process is repeated until the required convergence is achieved.

## 3. Monotonic convergence

The monotonic convergence of the algorithm is analyzed defining  $\Delta$  as the difference between the functional values before and after one iteration,

$$\begin{aligned} \Delta &\equiv J[\varphi^{(0)}, \varphi^{\dagger(0)}, \epsilon^{(0)}] - J[\varphi^{(1)}, \varphi^{\dagger(1)}, \epsilon^{(1)}] \\ &= \Delta_1 + \int_0^T (\Delta_{2a}(t) + \Delta_{2b}(t)) dt. \end{aligned} \quad (14)$$

The terms  $\Delta_j$  are derived in the Appendix and can be evaluated using Eqs. (4-8). This yields

$$\Delta_1 = -\lambda_0 \langle\zeta(T)|\hat{\mathbf{D}}|\zeta(T)\rangle, \quad (15)$$

cf. Eq. (A24), with the definition

$$|\zeta(t)\rangle = |\varphi^{(1)}(t)\rangle - |\varphi^{(0)}(t)\rangle. \quad (16)$$

Furthermore,

$$\begin{aligned} \Delta_{2a}(t) &= -g_a[\epsilon^{(1)}] + g_a[\epsilon^{(0)}] + \\ &\quad \left[\frac{\partial g_a}{\partial \epsilon}\right]_{(1)} (\epsilon^{(1)} - \epsilon^{(0)}), \end{aligned} \quad (17)$$

cf. Eq. (A29), which yields for our choice  $\epsilon_r = \epsilon^{(0)}$

$$\Delta_{2a}(t) = \lambda_a(t) (\epsilon^{(1)}(t) - \epsilon^{(0)}(t))^2, \quad (18)$$

and

$$\Delta_{2b}(t) = -\lambda_b \langle\zeta(t)|\hat{\mathbf{P}}(t)|\zeta(t)\rangle, \quad (19)$$

cf. Eq. (A29).

The algorithm converges monotonically to a minimum (maximum) of the functional if  $\Delta \geq 0$  ( $\Delta \leq 0$ ) in each iteration step. A sufficient but not necessary condition consists in all  $\Delta_j$  being larger (smaller) than zero. Let the operators  $\hat{\mathbf{D}}$  and  $\hat{\mathbf{P}}(t)$  be positive-semidefinite. Sufficient conditions are then given by

$$\lambda_0 \leq 0, \quad \lambda_b \leq 0, \quad \lambda_a(t) \geq 0. \quad (20)$$

for minimization, and by

$$\lambda_0 \geq 0, \quad \lambda_b \geq 0, \quad \lambda_a(t) \leq 0 \quad (21)$$

for maximization.

This result leads to some curious consequences. For example, let the system be described by a discrete number of levels and assume its Hilbert space can be split into two subspaces, the “allowed” subspace, described by the projector  $\hat{\mathbf{P}}_{allow}$ , and the “forbidden” subspace, described by  $\hat{\mathbf{P}}_{forbid}$  ( $\hat{\mathbf{P}}_{allow} + \hat{\mathbf{P}}_{forbid} = \hat{\mathbf{I}}$ ). The objective of the optimization consists in some transition inside the allowed subspace, avoiding any population transfer to the forbidden one. In the case of minimization of the functional  $J$ , the latter requirement can be expressed by one of the two following choices for  $J_b$ ,

$$\begin{aligned} (a) \quad \hat{\mathbf{P}}(t) &= \hat{\mathbf{P}}_{allow}, \quad \lambda_b \leq 0, \\ (b) \quad \hat{\mathbf{P}}(t) &= \hat{\mathbf{P}}_{forbid}, \quad \lambda_b \geq 0. \end{aligned} \quad (22)$$

In the case of maximization, the possibilities are

$$\begin{aligned} (a) \quad \hat{\mathbf{P}}(t) &= \hat{\mathbf{P}}_{allow}, \quad \lambda_b \geq 0, \\ (b) \quad \hat{\mathbf{P}}(t) &= \hat{\mathbf{P}}_{forbid}, \quad \lambda_b \leq 0. \end{aligned} \quad (23)$$

In both cases, (a) and (b) have the same physical meaning, remaining in the allowed subspace, or equivalently,

avoiding the forbidden subspace. The choice (b) is more appealing in principle, since the inhomogeneous term of Eq. (10) would decrease and eventually become negligible when approaching an optimal solution. However, only (a) fulfills the sufficient conditions for monotonic convergence.

A note of caution must be made at this point. Eqs. (20) and (21) are sufficient but not necessary conditions. Monotonic convergence can therefore be found for values of  $\lambda$  not fulfilling Eqs. (20) and (21). This can happen if the values of  $\Delta_j$  compensate each other to give a convergent total  $\Delta$ . In addition, the analysis assumes an exact solution of the control equations. A limited accuracy of the numerical implementation of the algorithm and a poor accuracy of the propagation method can lead to the breakdown of the monotonic convergence [26].

### B. Optimization of a unitary transformation

The objective consists in implementing a given unitary transformation  $\hat{\mathbf{O}}$ , up to a global phase, in a given subspace  $\mathcal{H}_{\hat{\mathbf{R}}}$  of dimension  $N_r$  described by the projector  $\hat{\mathbf{R}}$ ,

$$\hat{\mathbf{R}} = \sum_{n=1}^{N_r} |n\rangle\langle n|. \quad (24)$$

To this end, the parameter  $\tau$  is defined,

$$\tau = \text{Tr} \left\{ \hat{\mathbf{O}}^\dagger \hat{\mathbf{U}}(T, 0; \epsilon) \hat{\mathbf{R}} \right\} = \sum_{n=1}^{N_r} \langle \varphi_{fn} | \varphi_n(T) \rangle, \quad (25)$$

where

$$\begin{aligned} \hat{\mathbf{O}}|n\rangle &= |\varphi_{fn}\rangle, \\ \hat{\mathbf{U}}(t, 0; \epsilon)|n\rangle &= |\varphi_n(t)\rangle. \end{aligned} \quad (26)$$

The modulus of  $\tau$  is equal to  $N_r$  when the target unitary transformation is implemented in the subspace  $\mathcal{H}_{\hat{\mathbf{R}}}$  by the field  $\epsilon$  [6].

The optimization problem is again formulated as a functional minimization (maximization). The final time term is now defined by

$$\begin{aligned} J_0[\{\varphi_{Tn}, \varphi_{Tn}^\dagger\}] &= \lambda_0 |\tau|^2 \\ &= \lambda_0 \sum_{n=1}^{N_r} \langle \varphi_{fn} | \varphi_n(T) \rangle \sum_{n'=1}^{N_r} \langle \varphi_{n'}(T) | \varphi_{fn'} \rangle, \end{aligned} \quad (27)$$

where  $\{\varphi_n, \varphi_n^\dagger\}$  denote the set of states  $|\varphi_n\rangle$  ( $n = 1, \dots, N_r$ ). Other choices of  $J_0$  are possible [6]. The intermediate-time state-dependent term takes the form,

$$\begin{aligned} g_b[\{\varphi_n, \varphi_n^\dagger\}] &= \lambda_b \text{Tr} \left\{ \hat{\mathbf{U}}(t, 0; \epsilon)^\dagger \hat{\mathbf{P}}(t) \hat{\mathbf{U}}(t, 0; \epsilon) \hat{\mathbf{R}} \right\} \\ &= \lambda_b \sum_{n=1}^{N_r} \langle \varphi_n(t) | \hat{\mathbf{P}}(t) | \varphi_n(t) \rangle. \end{aligned} \quad (28)$$

The constraint over the field is taken to be the same as in the state-to-state case, cf. Eq. (8).

The equations defining the algorithm are obtained using the Krotov method as outlined in the Appendix. They read as follows:  $N_r$  “states”  $|\chi\rangle$  are given by the inhomogeneous evolution equation,

$$\frac{d}{dt} |\chi_n(t)\rangle = -\frac{i}{\hbar} \hat{\mathbf{H}}[\epsilon^{(0)}(t)] |\chi_n(t)\rangle + \lambda_b \hat{\mathbf{P}} |\varphi_n^{(0)}(t)\rangle, \quad (29)$$

with the “initial” condition

$$\begin{aligned} |\chi_n(T)\rangle &= -\lambda_0 \text{Tr} \left\{ \hat{\mathbf{O}}^\dagger \hat{\mathbf{U}}(T, 0; \epsilon^{(0)}) \hat{\mathbf{R}} \right\} |\varphi_{fn}\rangle \\ &= -\lambda_0 \left( \sum_{n'=1}^{N_r} \langle \varphi_{fn'} | \varphi_{n'}^{(0)}(T) \rangle \right) |\varphi_{fn}\rangle, \end{aligned} \quad (30)$$

$n = 1, \dots, N_r$ . They are used to determine the field  $\epsilon^{(1)}$  by means of

$$\begin{aligned} \epsilon^{(1)}(t) &= \\ \epsilon^{(0)}(t) &- \frac{1}{\hbar \lambda_a(t)} \text{Im} \left\{ \sum_{n=1}^{N_r} \langle \chi_n(t) | \hat{\mu} | \varphi_n^{(1)}(t) \rangle \right\}. \end{aligned} \quad (31)$$

The discussion of the monotony of convergence is equivalent to the state-to-state case [6]. That is, the sufficient conditions for minimization, Eq. (20), and maximization, Eq. (21) are also valid for the optimization of a unitary transformation.

### III. ILLUSTRATION OF THE ALGORITHM ENFORCING A STATE-DEPENDENT CONSTRAINT

In order to illustrate the algorithm outlined in the previous section, a simplified model of the vibrations in a  $\text{Rb}_2$  molecule is employed where only three electronic states are considered, cf. fig. 1. Of each electronic state, 11 vibrational levels were chosen, specifically  $v = 0, \dots, 10$  from the  $X^1\Sigma_g^+$  electronic ground state,  $v' = 5, \dots, 15$  from the  $^1\Sigma_u^+$  excited state, and  $v'' = 2, \dots, 12$  from the  $^1\Pi_g$  excited state. A laser field,  $\epsilon(t)$ , couples the  $X^1\Sigma_g^+$  levels to the  $^1\Sigma_u^+$  levels and the  $^1\Sigma_u^+$  to the  $^1\Pi_g$  levels. In this model the transitions  $v = 0$  to  $v' = 10$  and  $v' = 10$  to  $v'' = 6$  have similar frequencies and Franck-Condon factors (FCF),  $\omega_{1 \rightarrow 10} = 0.0507$  a.u. vs.  $\omega_{10 \rightarrow 6} = 0.0506$  a.u., and modulus of FCF 0.17 vs. 0.23. The Hamiltonian describing our model is then given by

$$\begin{aligned} \hat{\mathbf{H}} &= \sum_{i=1}^3 \hat{\mathbf{H}}_i \otimes |e_i\rangle\langle e_i| + \\ &\hat{\mu} \epsilon(t) (|e_1\rangle\langle e_2| + |e_2\rangle\langle e_1| + |e_2\rangle\langle e_3| + |e_3\rangle\langle e_2|), \end{aligned} \quad (32)$$

where a dipole moment operator  $\hat{\mu}$  independent of the internuclear distance  $R$  is assumed. The vibrational Hamiltonian is denoted by  $\hat{\mathbf{H}}_i$  and the electronic state associated to  $X^1\Sigma_g^+$ ,  $^1\Sigma_u^+$  and  $^1\Pi_g$  by  $|e_i\rangle$ , ( $i = 1, 2, 3$ ), respectively. The vibrational level energies and FCF were

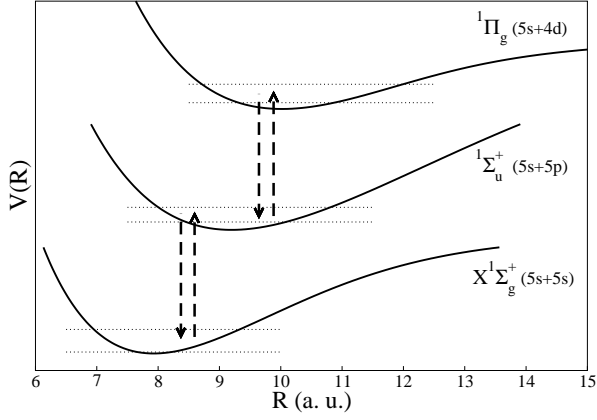


FIG. 1: Schematic representation of the  $\text{Rb}_2$  model used in the calculations. The dotted lines indicate the position of the vibrational manifolds considered in each electronic state.

obtained by diagonalization of the vibrational Hamiltonians employing the potential energy curves of Ref. [29].

For the optimization examples described below, a guess field of the form

$$\epsilon_g(t) = \epsilon_0 s(t) \cos[\Omega(t - T)/2] \quad (33)$$

is employed.  $T$  corresponds to the target time for the optimization, set to  $T = 8$  ps, and  $s(t) = \exp[-32(t/T - 1/2)^2]$  is taken to be a Gaussian shape function. The central frequency of the guess field is chosen to be  $\Omega = \omega_{v=0 \rightarrow v'=10}$  and  $\epsilon_0 = 10^{-4}$  a.u.. The optimization parameters related to the final time objective,  $J_0$  and to the field constraint,  $J_a$ , are set to  $\lambda_0 = -1$  and  $\lambda_a(t) = 100/s(t)$  in all calculations.

The state-dependent constraint in the optimization forces the population to remain in the subspace of the two lower electronic states. This is formulated by identifying  $\hat{\mathbf{P}}(t) \equiv \hat{\mathbf{P}}_{\text{allow}}$  as the projector onto the  $X^1\Sigma_g^+$  and  $1\Sigma_u^+$  levels,  $\hat{\mathbf{P}}_{\text{forbid}}$  corresponds thus to the projector onto the  $1\Pi_g$  levels. This choice of the allowed subspace is motivated as follows. The pulse duration is much shorter than the spontaneous emission lifetimes. On the timescale of the pulse, losses in a molecular system are due to processes such as predissociation, auto- or multiphoton ionization. Unlike spontaneous emission, these processes are relevant only in certain excited electronic states. Electronic states which are not affected by loss can therefore be included in the allowed subspace.

#### A. State-to-state optimization under state-dependent constraints

The objective for state-to-state optimization is chosen to transfer population initially in level  $v = 0$  of the electronic ground state to level  $v = 1$  at time  $T$ , using Raman-like transitions via levels  $v'$  in the  $1\Sigma_u^+$  excited state. Optimizations with and without state-dependent

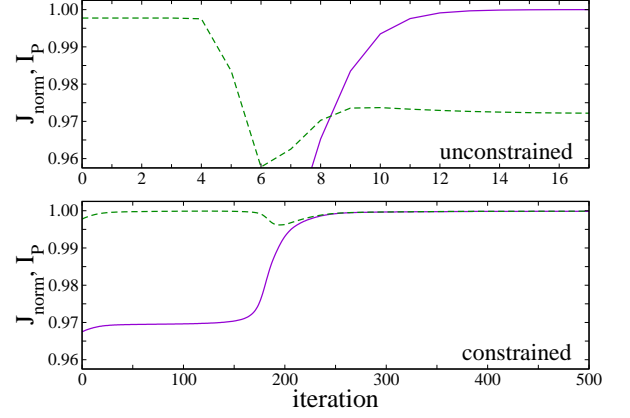


FIG. 2: (Color online)  $J_{\text{norm}}$  (violet solid line) and  $I_P$  (green dashed line) as a function of the number of iterations for  $\lambda_b = 0$  corresponding to optimization without state-dependent constraint (upper panel) and  $\lambda_b T = -32$  (lower panel). Note the different scale of the  $x$ -axes in the two panels.

constraint are compared. In both cases, the optimal field achieves a population transfer larger than 99.9% at the final time  $T$ . However, without the state-dependent constraint the optimal field does not “know” that it is not supposed to populate levels in the upper electronic  $1\Pi_g$  state at intermediate times. Due to similar transition frequencies and Franck-Condon factors, population transfer into the forbidden subspace therefore occurs.

In order to compare the performance of optimization with and without state-dependent constraint quantitatively, two measures are defined,

$$J_{\text{norm}} = \frac{J}{\lambda_0 + \lambda_b T}, \quad (34)$$

and

$$I_P = \frac{J_b}{\lambda_b T} = \frac{1}{T} \int_0^T \langle \varphi(t) | \hat{\mathbf{P}}_{\text{allow}} | \varphi(t) \rangle dt. \quad (35)$$

Unlike the original functional Eq. (9), the normalized functional  $J_{\text{norm}}$  has an optimal value equal to one independent of the choice of  $\lambda_0$  and  $\lambda_b$ , while  $I_P$  corresponds to the average value of population in the allowed subspace. Figure 2 shows the values of  $J_{\text{norm}}$  and  $I_P$  as the iterative optimization proceeds. Optimization with state-dependent constraint requires a larger number of iterations to reach the optimal value of the total objective  $J$ . However, this price is paid off since in this case the solution keeps indeed nearly all of the population in the allowed subspace at any time.

Figures 3 and 4 demonstrate the behavior of the populations under the dynamics generated by the optimized pulses obtained with and without the state-dependent constraint. The amount of population in the upper electronic state is largely reduced for a field obtained with state-dependent constraint as compared to that resulting from unconstrained optimization, cf. fig. 3. Two different transfer mechanisms are found: The pulse obtained

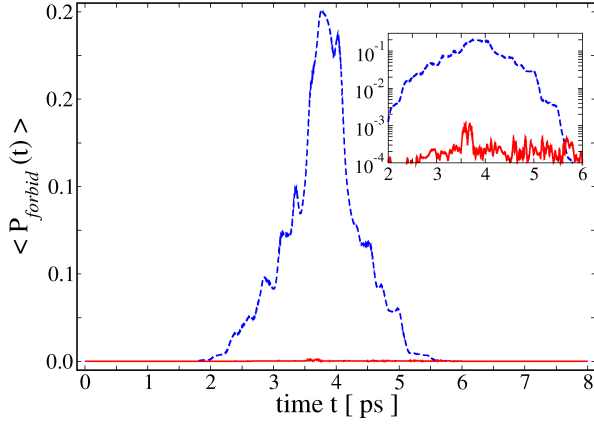


FIG. 3: (Color online) Population in the forbidden subspace as a function of time for the system driven by the optimal field obtained with  $\lambda_b = 0$  after 17 iterations (blue dashed line) and with  $\lambda_b T = -32$  after 500 iterations (red solid line). The inset shows the population in the central time interval in a semi-logarithmic plot.

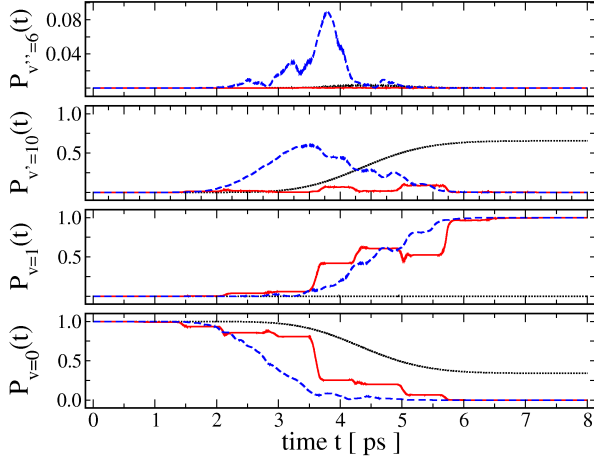


FIG. 4: (Color online) Population in levels  $v = 0$ ,  $v = 1$ ,  $v' = 10$  (in the allowed subspace) and  $v'' = 6$  (in the forbidden subspace) as a function of time for the system driven by the optimal field obtained with  $\lambda_b = 0$  after 17 iterations (blue dashed line) and with  $\lambda_b T = -32$  after 500 iterations (red solid line). Also shown is the population for the evolution with the guess field (black dotted line). Note the different scale of the  $y$ -axis for  $v'' = 6$ .

with the state-dependent constraint transfers population to  $v = 1$  in a ladder-like process which is driven by short subpulses. In between the subpulses, the amount of population in the  $^1\Sigma_u^+$  excited state (level  $v' = 10$ ) is small, cf. fig. 4. Further excitation to the forbidden  $^1\Pi_g$  levels becomes thus unlikely. On the other hand, the pulse obtained without the state-dependent constraint transfers a large amount of population to the intermediate  $^1\Sigma_u^+$  electronic state (mainly to level  $v' = 10$ ) which is later to be brought back to the ground state to level  $v = 1$ . The large amount of population which resides in the  $^1\Sigma_u^+$  electronic state while the field is on allows for transient transfer to

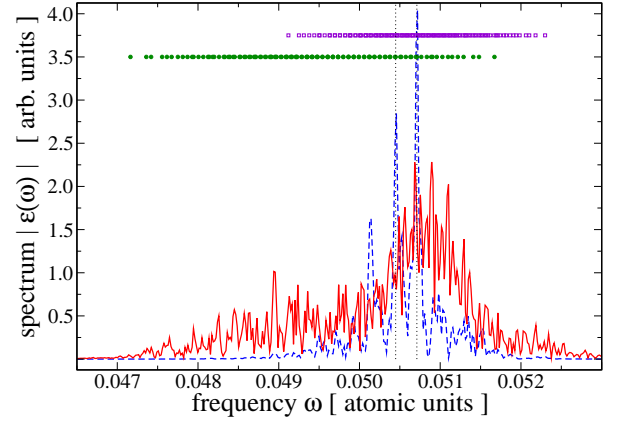


FIG. 5: (Color online) Spectral amplitude as a function of the frequency  $\omega$  for of the optimized field obtained with  $\lambda_b = 0$  after 17 iterations (blue dashed line) and with  $\lambda_b T = -32$  after 500 iterations (red solid line). The dashed vertical lines indicate the transition frequencies  $\omega_{v=0 \rightarrow v'=10}$  and  $\omega_{v=1 \rightarrow v'=10}$ . The transition frequencies between the levels of the  $X^1\Sigma_g^+$  electronic ground state and the  $^1\Sigma_u^+$  intermediate state are represented by green circles, and the transition frequencies between the  $^1\Sigma_u^+$  intermediate state and the  $^1\Pi_g$  upper state by violet squares.

the upper electronic state at intermediate times.

Both transfer mechanisms share the common feature of a process driven mainly by one-photon transitions. Population is transferred to the intermediate electronic state by a one-photon absorption. This population is later sent back to the electronic ground state by a one-photon emission or further excited to the upper electronic state by another one-photon absorption. Large spectral amplitudes of the optimal fields at frequencies corresponding to the main transition frequencies of the system reflect this finding, cf. fig. 5. Analysis of the two-photon spectrum, i.e. of the Fourier transform of  $\epsilon(t)\epsilon(t)$ , confirms that the amount of processes involving two or more photons is very small. Figure 5 furthermore illustrates an important difference between the results of optimization with and without state-dependent constraint: The additional requirement implies a more complex optimal solution which is reflected both in a broader spectrum and in a more intricate dependence of the spectrum on frequency.

## B. Optimization of a unitary transformation under state-dependent constraints

The implementation of a Fourier transform [30] in levels  $v = 0, 1, 2, 3$  of the  $X^1\Sigma_g^+$  electronic ground state is chosen as example objective for the optimization of a unitary transformation. The state-dependent constraint is taken to be identical to the optimization of state-to-state transfer, i.e. the  $^1\Pi_g$  upper electronic state represents the forbidden subspace. Figure 6 shows the population

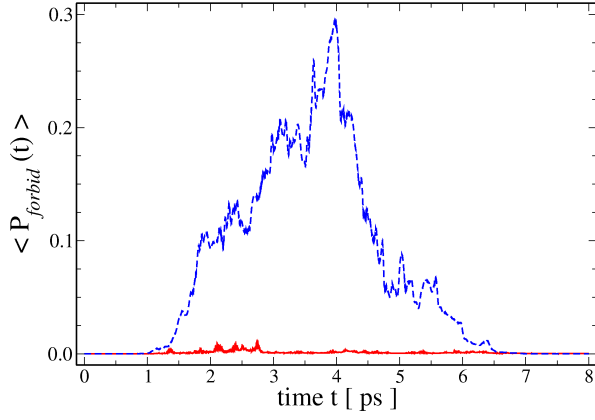


FIG. 6: (Color online) Population in the forbidden subspace as a function of time for the system driven by the optimal field obtained with  $\lambda_b = 0$  after 50 iterations (blue dashed line) and with  $\lambda_b T = -8$  after 500 iterations (red solid line) when the system is initially in level  $v = 0$ .

in the levels of the forbidden subspace for the system driven by an optimized field obtained with and without state-dependent constraint. In both cases  $|\tau| > 3.999$ ; since the target value is equal to 4, the number of levels in which the unitary transformation is implemented, this corresponds to an error of less than  $10^{-3}$ . The main results are similar to those of optimizing a state-to-state transition: A larger number of iterations is needed to obtain similar performance with respect to the final time objective  $J_0$ . In addition the solution becomes more complex for the optimization with the state-dependent constraint. The two tasks of a state-to-state transition and of a unitary transformation are run with the same number of iterations for optimization with the state-dependent constraint. As expected the state-to-state transition converges faster, cf. figs. 3 and 6. The effort for the optimization of a unitary transformation is approximately equivalent to  $N$  simultaneous state-to-state transitions (here  $N = 4$ ), i.e. it corresponds to a more difficult optimization problem [6].

### C. Robustness with respect to decay

A loss mechanism in the  $^1\Pi_g$  electronic state is modeled by adding an imaginary term  $-i\Gamma/2$  to the vibrational energies, where  $\Gamma = 1/\tau_L$  denotes the decay rate and  $\tau_L$  the lifetime. Physically, such a decay is caused by processes such as predissociation or auto-ionization. When decay is included, the system dynamics generated by the optimized field is perturbed depending on the decay rate and on the amount of population in the lossy upper electronic state. Figure 7a shows the population in the target level for the state-to-state transition with and without state-dependent constraint with optimization parameters as described above (solid lines). Moreover, the total amount of population remaining in the

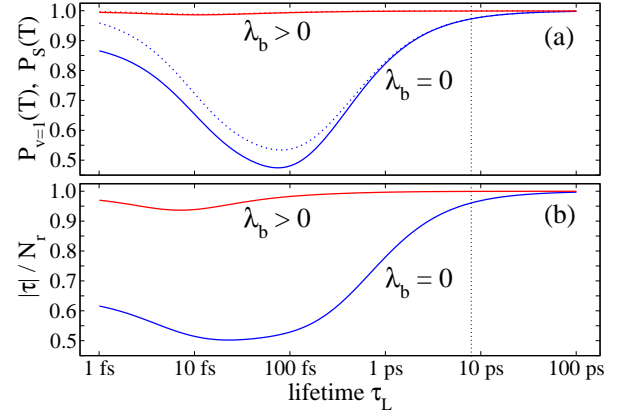


FIG. 7: (Color online) Comparison of final results at time  $t = T$  as a function of the lifetime of the upper excited state. (a) State-to-state transition: Population in level  $v = 1$ ,  $P_{v=1}(T)$ , (solid lines) and total population remaining in the system,  $P_s(T)$ , (dotted lines) for the optimized field obtained with (red) and without the state-dependent constraint (blue). (b) Unitary transformation: Normalized final time objective,  $|\tau|/N_r$  for a field obtained by optimization of the unitary transformation with (red solid line) and without state-dependent constraint (blue solid line). The vertical dotted line indicates the overall pulse duration.

system at time  $T$ ,

$$P_s(T) = \langle \varphi(T) | (\hat{\mathbf{P}}_{allow} + \hat{\mathbf{P}}_{forbid}) | \varphi(T) \rangle, \quad (36)$$

is depicted by dotted lines in fig. 7a. The smaller population transfer to the upper state in case of optimization with the state-dependent constraint results in a larger robustness of the solution as the decay rate is increased. For a lifetime on the order of the pulse duration, the final time objective is only very slightly perturbed. In contrast the example without state-dependent constraint shows already a significant loss in the objective. This effect becomes even more evident when the lifetime is smaller than the pulse duration. For optimization without state-dependent constraint the transfer efficiency is reduced by 50% for lifetimes on the order of 100 fs. For short lifetimes both optimization methods fail, nevertheless the algorithm including the state-dependent constraint is still superior. The decrease in efficiency of the transfer to the  $v = 1$  target level (solid lines) is associated to the loss of population from the system (dotted lines) as the lifetime decreases.

An interesting effect is found as the lifetime becomes very small: the final time objective actually improves. This is a manifestation of the quantum Zeno effect which states that a continuously observed quantum system never decays [31]. In our example, the usual interpretation is inverted: The decay process can be associated with a weak measurement monitoring the population in the third electronic state, and the optimized field corresponds to the decay in the usual picture. As the decay rate increases, respectively the lifetime decreases, monitoring of the forbidden subspace becomes continuous and

the pulse cannot populate the lossy state anymore. The possibility to loose population becomes thus smaller (cf. dotted lines in fig. 7a) which entails also better results for the final time objective (cf. solid lines in fig. 7a). However, by comparing the solid and dotted lines in fig. 7a, it is obvious that the loss of coherence in the control process is larger than the loss of population alone.

Loss or decoherence is the main obstacle for generating controlled unitary transformations. Figure 7b shows how the objective deteriorates due to loss for the unitary transformation. The optimized field obtained with the state-dependent constraint is able to maintain a high fidelity despite the loss term. The performance of the optimized field for the unitary transformation is clearly worse than for the state-to-state transition. This is due to more population transfer to the upper electronic state in case of the unitary transformation obtained after the same number of iterations. A result comparable to that of optimizing the state-to-state transition with the state-dependent constraint can be achieved by increasing the number of iterations.

#### IV. DISCUSSION

The present work is related to a number of previous OCT studies using state-dependent constraints, or equivalently, a time dependent target. The general formulation of our complete functional for optimization of a state-to-state transition employing a state-dependent constraint, Eq. (9), is related to the functional of Ref. [24] identifying  $\hat{\mathbf{D}} = X$  and  $\hat{\mathbf{P}}(t) = Y(t)$ . Similarly, the specific application of Refs. [25, 27] corresponds to  $J_0 = 0$  and  $\hat{\mathbf{P}}(t) = f(t)\hat{\mathbf{O}}$ ; and the application of Ref. [26] to  $J_0 = 0$ ,  $\hat{\mathbf{P}}(t) = \hat{\mathbf{O}}_1(t) + 2T\delta(t-T)\hat{\mathbf{O}}_2$ , and  $\hat{\mathbf{D}} = \hat{\mathbf{O}}_2$ . The approaches of Refs. [25, 26, 27] consist in imposing a specific desired dynamics onto the system, for example, the populations following a given time profile, or the maximization of the expectation value of a given operator.

In the studies of Refs. [24, 25, 26, 27], the optimization algorithm was obtained using the variational method. It is well-known that the variational approach is compatible with a large number of implementations of the control algorithm, not all of them showing monotonic convergence. Ref. [24] gives a detailed analysis of this family. Our results which were obtained with the Krotov method, correspond to the case  $\eta_k = 0$  and  $\xi_k = 1$  of Ref. [24]. The Krotov method comes with the advantage of allowing a straightforward discussion of the convergence of the algorithm in terms of the sign of the optimization parameters  $\lambda_i$ . As discussed in Section II, this sign of the optimization parameter turns out to be crucial for the choice of projector  $\hat{\mathbf{P}}(t)$ .

As mentioned in the introduction, a different approach to avoid population leakage to undesired states is using local control theory [19, 20, 21, 22]. The example chosen in Ref. [22] differs from ours: In Ref. [22], the allowed

subspace was restricted to the register levels of a unitary transformation. This choice forces the dynamics of the system to rely solely on two-photon processes which assure that the population remains in the register. A similar task in our formulation would correspond to choosing the operator  $\hat{\mathbf{P}}_{allow}$  as the projector onto the  $X^1\Sigma_g^+$  levels only. However, the state-dependent constraint is formulated such as to maximize the time-averaged population in the allowed subspace. The family of solutions found by the algorithm consists in transferring some amount of population to the intermediate electronic state, but only for a very short time. This is reflected in the inset of fig. 3 in a sequence of “spikes” in the population of the forbidden subspace as a function of time. Since the constraint of not populating any electronically excited state is much more restrictive than the time-averaged formulation, the number of optimization steps needed for reaching a specified efficiency and hence the complexity of the solution increase largely compared to the present results.

#### V. CONCLUSIONS

Steering a system to a desired objective has to always be balanced by the damage induced by the steering process. In the present study optimal control theory was adopted to include a positive objective to be maximized and a negative constraint to be avoided. To ensure monotonic convergence the additional constraint which depends on the state of the system was incorporated in the Krotov method.

It turned out that the state-dependent constraint needs to be formulated in terms of maximizing population in the allowed subspace. While one could expect this to be equivalent to minimizing population in the forbidden subspace, the resulting algorithms and their convergence behavior differ markedly.

The algorithm was applied to a simple model mimicking vibrational manifolds in three electronic states of an  $\text{Rb}_2$  molecule. Population transfer from the vibrational level  $v = 0$  to  $v = 1$  of the electronic ground state and the implementation of a Fourier transform in the levels  $v = 0, 1, 2, 3$  were chosen as optimization examples for a state-to-state transition and for a unitary transformation. In both cases, the optimized fields induce Raman-like transitions via an electronically excited state. It was shown that optimization including the state-dependent constraint indeed avoids population of a higher electronically excited state. This state can either correspond to a loss channel itself or represent a resonant intermediate state in an unwanted multi-photon process.

A similar task of promoting one objective while avoiding damage has been developed using local control theory [20]. It is important to note that the optimal solutions are quite different. The local control scheme tries to minimize the instantaneous population. Therefore it resorts to off-resonant two-photon transitions where only a transient population exists. The OCT scheme minimizes the



integrated population in the forbidden subspace. As a result the solution can contain abrupt spikes of population. Since these spikes have no time to go anywhere the system remains protected against damage.

The success of the scheme was demonstrated in the ability to cope with a real decay channel in the forbidden subspace. The decay term causes loss of population and loss of phase or decoherence. The state-to-state objective can be thought to work without long term coherence. Brief periods of coherence are sufficient to generate the transitions. On the contrary the unitary operator objective has to maintain phase coherence for the total period. The ability of the algorithm to find solutions that can cope with this scenario is encouraging. The robustness of the coherent control solution could result from a scheme based on a large number of interfering pathways. In this case loss of a few pathways will only slightly hinder the final objective.

### Acknowledgments

We would like to thank David Tannor for many fruitful discussions. We gratefully acknowledge financial support from the Spanish MCT (FIS2004-05687, FIS2007-64018) and the Gobierno de Canarias (PI2004/025), from the Israel Science Foundation, and from the Deutsche Forschungsgemeinschaft (KO 2302/1-1). The Fritz Haber Center is supported by the Minerva Gesellschaft für die Forschung GmbH München, Germany.

## APPENDIX A: REVIEW OF THE KROTOV METHOD FOR OPTIMAL CONTROL THEORY

In the following the system Hamiltonian  $\hat{H}$  is assumed to be Hermitian. The derivation can easily be generalized to non-Hermitian Hamiltonians [24]. Let's define the states  $|f\rangle$ ,

$$\begin{aligned} |f\rangle &= -\frac{i}{\hbar}\hat{H}[\epsilon]|\varphi\rangle, \\ \langle f| &= \langle\varphi|\frac{i}{\hbar}\hat{H}[\epsilon], \end{aligned} \quad (\text{A1})$$

which correspond to the total time derivative of  $\varphi(t)$  in Eq. (1). As mentioned above, the rigorous utilization of the Krotov method is somewhat cumbersome and little instructive. For simplicity the case of a functional depending on two real functions denoted by  $\varphi$  and  $\varphi^\dagger$  and a real control  $\epsilon$  will be presented. This simple case can be connected to the original problem by the following relations,

$$\begin{aligned} \psi &\longleftrightarrow |\psi\rangle, \\ \psi^\dagger &\longleftrightarrow \langle\psi|. \end{aligned} \quad (\text{A2})$$

for  $\psi = \varphi, \chi, f$ . Our derivation follows closely Ref. [5], but including the state-dependent constraint, i.e. the dependency on  $g_b$ , cf. Eq. 6. The final equations for the original problem, Eqs. (10-12), are obtained following the steps presented in this Appendix. Moreover, our outline can be employed to derive new optimization schemes based on different choices of  $J_0$ ,  $g_a$  or  $g_b$ .

### 1. The scalar function $\Phi$ and the functional $L$

All functions considered in the derivation,  $\varphi$ ,  $\epsilon$ , etc., depend on  $t$  but for simplicity this dependence will only be made explicit for the initial and the final time, e.g.  $\varphi_0$ ,  $\varphi_T$ . The Krotov method is based on the introduction of the arbitrary scalar function  $\Phi(t, \varphi, \varphi^\dagger)$ , the functions

$$G(\varphi_T, \varphi_T^\dagger, \Phi_T) = J_0(\varphi_T, \varphi_T^\dagger) + \Phi(T, \varphi_T, \varphi_T^\dagger), \quad (\text{A3})$$

and

$$R(\varphi, \varphi^\dagger, \epsilon, \Phi) = -g_a(\epsilon) - g_b(\varphi, \varphi^\dagger) + \left[ \frac{\partial \Phi}{\partial \varphi} \right] f + f^\dagger \left[ \frac{\partial \Phi}{\partial \varphi^\dagger} \right] + \frac{\partial \Phi}{\partial t}. \quad (\text{A4})$$

Note that

$$R = -(g_a + g_b) + \frac{d\Phi}{dt}, \quad (\text{A5})$$

where it was used that  $f$  ( $f^\dagger$ ) is the total time derivative of  $\varphi$  ( $\varphi^\dagger$ ).

A new functional can be defined

$$L[\varphi, \varphi^\dagger, \epsilon, \Phi] = G(\varphi_T, \varphi_T^\dagger, \Phi_T) - \Phi(0, \varphi_0, \varphi_0^\dagger) - \int_0^T R(\varphi, \varphi^\dagger, \epsilon, \Phi) dt, \quad (\text{A6})$$

with the interesting property

$$L[\varphi, \varphi^\dagger, \epsilon, \Phi] = J[\varphi, \varphi^\dagger, \epsilon]. \quad (\text{A7})$$

The Krotov method takes advantage of this property and the freedom in the choice of  $\Phi$  to find an iterative algorithm that minimizes (maximizes) the original functional  $J$  [5, 6].

## 2. The iterative algorithm

Let's start with a given field  $\epsilon^{(0)}$  and the corresponding functions  $\varphi^{(0)}$  and  $\varphi^{\dagger(0)}$ . The values of the functionals are denoted by  $J^{(0)}$  and  $L^{(0)}$ . The functional  $J$  ( $L$ ) should be minimized (maximization will be discussed later on). The objective is then to determine a new control  $\epsilon^{(1)}$ , given the functions  $\varphi^{(1)}$  and  $\varphi^{\dagger(1)}$ , for which

$$J^{(0)} = L^{(0)} \geq J^{(1)} = L^{(1)}. \quad (\text{A8})$$

For the time being, the control  $\epsilon$  and the functions  $\varphi$ ,  $\varphi^\dagger$  are treated as “independent variables”. The Krotov method accomplishes the optimization in two steps.

1. A function  $\Phi$  is determined such that  $L[\varphi, \varphi^\dagger, \epsilon, \Phi]$  has a maximum for  $\varphi^{(0)}$  y  $\varphi^{\dagger(0)}$  regardless of the control  $\epsilon$ . The maximum condition implies second order (functional) derivatives of  $L$  with respect to  $\varphi$  and  $\varphi^\dagger$ . Note that these derivatives should be evaluated “at” any field  $\epsilon$ . However, the much simpler evaluation at  $\epsilon^{(0)}$  turns out to be sufficient [5].
2.  $\epsilon^{(1)}$  is determined such that the functional is minimized with respect to all possible controls  $\epsilon$ . This implies a second order (functional) derivative of  $L$  with respect to  $\epsilon$ . Since  $\Phi$  is determined by the conditions of step 1, this can be done without considering the effect of the change in the functional due to change in  $\varphi$  and  $\varphi^\dagger$  induced by the new control.

Finally, the necessary relation between the field and the states through the evolution equation is imposed, resulting in

$$L[\varphi^{(0)}, \varphi^{\dagger(0)}, \epsilon^{(0)}, \Phi] \geq L[\varphi^{(1)}, \varphi^{\dagger(1)}, \epsilon^{(1)}, \Phi]. \quad (\text{A9})$$

The conditions derived from the second order (functional) derivatives are rather complicated [5]. Fortunately, for the problems under consideration, the minimum and maximum conditions can be relaxed to extremum conditions on the functional, for which only first order (functional) derivatives are needed. Moreover, the

extremum conditions are common for minimization and maximization of the functional, therefore both cases are treated together below. Due to the relaxation of the minimum and maximum to extremum conditions, an additional step is required: The monotonic convergence to the target value of the final algorithm must be checked.

## 3. First step: Determining $\Phi$ up to first order

The evaluation of an expression [...] at  $\epsilon^{(0)}$ ,  $\varphi^{(0)}$  and  $\varphi^{\dagger(0)}$  is denoted by [...]  $_{(0)}$ . The extremum in  $L$  for  $\varphi^{(0)}$  and  $\varphi^{\dagger(0)}$  corresponds to the following conditions on  $R$ :

$$\left[ \frac{\partial R}{\partial \varphi} \right]_{(0)} = 0, \quad \left[ \frac{\partial R}{\partial \varphi^\dagger} \right]_{(0)} = 0. \quad (\text{A10})$$

For example, using Eq. (A4), the second derivative reads

$$\left[ \frac{\partial R}{\partial \varphi^\dagger} \right] = -\frac{\partial g_b}{\partial \varphi^\dagger} + \left[ \frac{\partial \Phi}{\partial \varphi} \right] \frac{\partial f}{\partial \varphi^\dagger} + \frac{\partial f^\dagger}{\partial \varphi^\dagger} \left[ \frac{\partial \Phi}{\partial \varphi^\dagger} \right] + \frac{d}{dt} \left[ \frac{\partial \Phi}{\partial \varphi^\dagger} \right], \quad (\text{A11})$$

where we have used that

$$\left[ \frac{\partial}{\partial \varphi} \frac{\partial \Phi}{\partial \varphi^\dagger} \right] f + f^\dagger \left[ \frac{\partial}{\partial \varphi^\dagger} \frac{\partial \Phi}{\partial \varphi} \right] + \frac{\partial}{\partial t} \left[ \frac{\partial \Phi}{\partial \varphi^\dagger} \right] = \frac{d}{dt} \left[ \frac{\partial \Phi}{\partial \varphi^\dagger} \right]. \quad (\text{A12})$$

New functions  $\chi$  and  $\chi^\dagger$  are defined,

$$\chi(t) \equiv \left[ \frac{\partial \Phi}{\partial \varphi^\dagger} \right]_{(0)}, \quad \chi^\dagger(t) \equiv \left[ \frac{\partial \Phi}{\partial \varphi} \right]_{(0)}. \quad (\text{A13})$$

The second condition in Eq. (A10) becomes then

$$\frac{d\chi}{dt} = -\chi^\dagger \left[ \frac{\partial f}{\partial \varphi^\dagger} \right]_{(0)} - \left[ \frac{\partial f^\dagger}{\partial \varphi^\dagger} \right]_{(0)} \chi + \left[ \frac{\partial g_b}{\partial \varphi^\dagger} \right]_{(0)}, \quad (\text{A14})$$

and a similar result can be found for  $\chi^\dagger$ . Moreover, the extremum condition on  $L$  also implies,

$$\left[ \frac{\partial G}{\partial \varphi_T} \right]_{(0)} = 0, \quad \left[ \frac{\partial G}{\partial \varphi_T^\dagger} \right]_{(0)} = 0. \quad (\text{A15})$$

For example the second derivative gives

$$\frac{\partial G}{\partial \varphi_T^\dagger} = \frac{\partial J_0}{\partial \varphi_T^\dagger} + \frac{\partial \Phi}{\partial \varphi_T^\dagger}. \quad (\text{A16})$$

Using the previous definition of  $\chi$ ,  $\chi(T)$  can be identified,

$$\chi(T) = \chi_T = - \left[ \frac{\partial J_0}{\partial \varphi_T^\dagger} \right]_{(0)}. \quad (\text{A17})$$

Therefore the extremum conditions allow to determine the function  $\chi$  (and  $\chi^\dagger$ ) using the evolution Eq. (A14) with the “initial” condition at time  $t = T$ , Eq. (A17). With the knowledge of  $\chi$  and  $\chi^\dagger$ , the function  $\Phi$  can be constructed up to first order,

$$\Phi_1[t, \varphi, \varphi^\dagger] = \chi^\dagger(t) \varphi + \varphi^\dagger \chi(t). \quad (\text{A18})$$

The function  $\Phi_1$  of the original problem is obtained using Eq. (A2),

$$\begin{aligned} \Phi_1[t, \varphi, \varphi^\dagger] &= \langle \chi(t) | \varphi(t) \rangle + \langle \varphi(t) | \chi(t) \rangle \\ &= 2\text{Re} \{ \langle \chi(t) | \varphi(t) \rangle \}. \end{aligned} \quad (\text{A19})$$

#### 4. Second step: Determining $\epsilon^{(1)}$

The extremum condition on  $L$  with respect to the control, evaluated in  $\epsilon^{(1)}$  leads to

$$\left[ \frac{\partial R}{\partial \epsilon} \right]_{(1)} = 0, \quad (\text{A20})$$

where  $[\dots]_{(1)}$  denotes the evaluation of the expression  $[\dots]$  at  $\epsilon^{(1)}$ ,  $\varphi^{(1)}$  and  $\varphi^{\dagger(1)}$ . Using the definition of  $R$ ,

---


$$L[\varphi^{(0)}, \varphi^{\dagger(0)}, \epsilon^{(0)}, \Phi_1^{(0)}] - L[\varphi^{(1)}, \varphi^{\dagger(1)}, \epsilon^{(1)}, \Phi_1^{(1)}] = \Delta_1 + \int_0^T \Delta_2(t) dt, \quad (\text{A23})$$


---

where

$$\begin{aligned} \Delta_1 &= G(\varphi_T^{(0)}, \varphi_T^{\dagger(0)}, \Phi_{1T}^{(0)}) - G(\varphi_T^{(1)}, \varphi_T^{\dagger(1)}, \Phi_{1T}^{(1)}) \\ &= J_0^{(0)} - \Phi_{1T}^{(0)} - J_0^{(1)} + \Phi_{1T}^{(1)}, \end{aligned} \quad (\text{A24})$$

and

$$\Delta_2(t) = R(\varphi^{(1)}, \varphi^{\dagger(1)}, \epsilon^{(1)}, \Phi_1^{(1)}) - R(\varphi^{(0)}, \varphi^{\dagger(0)}, \epsilon^{(0)}, \Phi_1^{(0)}). \quad (\text{A25})$$


---

the derivative gives

$$\frac{\partial R}{\partial \epsilon} = -\frac{\partial g_a}{\partial \epsilon} + \left[ \frac{\partial \Phi}{\partial \varphi} \right] \frac{\partial f}{\partial \epsilon} + \frac{\partial f^\dagger}{\partial \epsilon} \left[ \frac{\partial \Phi}{\partial \varphi^\dagger} \right], \quad (\text{A21})$$

where it was used that  $\Phi$  does not explicitly depend on  $\epsilon$ . Using  $\Phi_1$ , the previous equation results in,

$$-\left[ \frac{\partial g_a}{\partial \epsilon} \right]_{(1)} + \chi^\dagger \left[ \frac{\partial f}{\partial \epsilon} \right]_{(1)} + \left[ \frac{\partial f^\dagger}{\partial \epsilon} \right]_{(1)} \chi = 0. \quad (\text{A22})$$

Since  $g_a$  and the functions  $f$  and  $f^\dagger$ , defined in Eq. (A1), depend on the control, this equation allows to determine  $\epsilon^{(1)}$ .

#### 5. Conditions for monotonic convergence

To check whether the iterative method shows monotonic convergence, we define  $\Delta_1$  and  $\Delta_2(t)$ ,

To simplify this expression, it is assumed that

$$\frac{\partial f^\dagger}{\partial \varphi} = 0, \quad \frac{\partial f}{\partial \varphi^\dagger} = 0, \quad (\text{A26})$$

which is true for the original problem, and the function  $R$  is split into  $R = R_a + R_b$ ,

---


$$R_a = -g_a - \chi^* \left[ \frac{\partial f}{\partial \varphi} \right]_{(0)} \varphi + \chi^\dagger f - \varphi^\dagger \left[ \frac{\partial f^\dagger}{\partial \varphi^\dagger} \right]_{(0)} \chi + f^\dagger \chi, \quad (\text{A27})$$

$$R_b = -g_b + \left[ \frac{\partial g_b}{\partial \varphi} \right]_{(0)} \varphi + \varphi^\dagger \left[ \frac{\partial g_b}{\partial \varphi^\dagger} \right]_{(0)}. \quad (\text{A28})$$


---

Therefore  $\Delta_2(t) = \Delta_{2a}(t) + \Delta_{2b}(t)$  with

$$\begin{aligned} \Delta_{2a}(t) &= [R_a]_{(1)} - [R_a]_{(0)}, \\ \Delta_{2b}(t) &= [R_b]_{(1)} - [R_b]_{(0)}. \end{aligned} \quad (\text{A29})$$

If the objective is to minimize the functional  $J$  ( $J^{(0)} \geq J^{(1)}$ ), the sufficient but not necessary conditions for monotonic convergence are given by  $\Delta_j \geq 0$ , ( $j = 1, 2, 3$ ).

Analogously, in order to maximize the functional it is sufficient that  $\Delta_j \leq 0$ , ( $j = 1, 2, 3$ ). Whether the  $\Delta_j$  are positive or negative is determined by the particular choice of  $J_0$ ,  $g_a$ , and  $g_b$ . In Sec. II the values of  $\Delta$  are analyzed for the cases under study.

## 6. The Krotov method for unitary transformations

The previous approach can easily be generalized to the case of unitary transformations. The functional  $J$  depends then on  $2N_r$  real functions denoted by the set  $\{\varphi_n, \varphi_n^\dagger\}$ . As a consequence, the scalar functions  $\Phi$ ,  $G$  and  $R$  will also depend on all of them. Moreover, the new

dependence must be taken into account in the derivatives of the previous equations by the substitution

$$\begin{aligned} \frac{\partial}{\partial \varphi} &\longrightarrow \sum_n \frac{\partial}{\partial \varphi_n}, \\ \frac{\partial}{\partial \varphi^\dagger} &\longrightarrow \sum_n \frac{\partial}{\partial \varphi_n^\dagger}. \end{aligned} \quad (\text{A30})$$

The remaining procedure is analogous to the state-to-state case, and the relations (A2) can be used to obtain the equations for the optimization algorithm given in Sec. II B.

- 
- [1] S. A. Rice and M. Zhao, *Optical control of molecular dynamics* (John Wiley & Sons, New York, 2000).
  - [2] P. Brumer and M. Shapiro, *Principles and Applications of the Quantum Control of Molecular Processes* (Wiley Interscience, 2003).
  - [3] W. Zhu, J. Botina, and H. Rabitz, *J. Chem. Phys.* **108**, 1953 (1998).
  - [4] R. Kosloff, S. A. Rice, P. Gaspard, S. Tersigni, and D. J. Tannor, *Chem. Phys.* **139**, 201 (1989).
  - [5] S. E. Sklarz and D. J. Tannor, *Phys. Rev. A* **66**, 053619 (2002).
  - [6] J. P. Palao and R. Kosloff, *Phys. Rev. A* **68**, 062308 (2003).
  - [7] C. Gollub, M. Kowalewski, and R. de Vivie-Riedle, *quant-ph/0801.3935* (2008).
  - [8] G. M. Huang, T. J. Tarn, and J. W. Clark, *J. Math. Phys.* **24**, 2608 (1983).
  - [9] V. Ramakrishna, M. V. Salapaka, M. Dahleh, H. Rabitz, and A. Peirce, *Phys. Rev. A* **51**, 960 (1995).
  - [10] H. Rabitz, T.-S. Ho, M. Hsieh, R. Kosut, and M. Demiralp, *Phys. Rev. A* **74**, 012721 (2006).
  - [11] A. Bartana, R. Kosloff, and D. J. Tannor, *J. Chem. Phys.* **106**, 1435 (1997).
  - [12] A. Bartana, R. Kosloff, and D. J. Tannor, *Chem. Phys.* **267**, 195 (2001).
  - [13] Y. Ohtsuki, W. Zhu, and H. Rabitz, *J. Chem. Phys.* **110**, 9825 (1999).
  - [14] J. P. Palao and R. Kosloff, *Phys. Rev. Lett.* **89**, 188301 (2002).
  - [15] C. Tesch and R. de Vivie-Riedle, *Phys. Rev. Lett.* **89**, 157901 (2002).
  - [16] H. A. Rabitz, M. M. Hsieh, and C. M. Rosenthal, *Science* **303**, 1998 (2004).
  - [17] S. Kallush and R. Kosloff, *Phys. Rev. A* **73**, 032324 (2006).
  - [18] D. Stefanatos, N. Khaneja, and S. J. Glaser, *Phys. Rev. A* **69**, 022319 (2004).
  - [19] R. Kosloff, A. D. Hammerich, and D. Tannor, *Phys. Rev. Lett.* **69**, 2172 (1992).
  - [20] A. Bartana, R. Kosloff, and D. J. Tannor, *J. Chem. Phys.* **99**, 196 (1993).
  - [21] V. S. Malinovsky, C. Meier, and D. J. Tannor, *Chem. Phys.* **221**, 67 (1997).
  - [22] S. E. Sklarz and D. J. Tannor, *quant-ph/0404081* (2004).
  - [23] S. Sklarz and D. J. Tannor, *Chem. Phys.* **322**, 87 (2006).
  - [24] Y. Ohtsuki, G. Turinici, and H. Rabitz, *J. Chem. Phys.* **120**, 5509 (2004).
  - [25] A. Kaiser and V. May, *J. Chem. Phys.* **121**, 2528 (2004).
  - [26] I. Šerban, J. Werschnik, and E. K. U. Gross, *Phys. Rev. A* **71**, 053810 (2005).
  - [27] A. Kaiser and V. May, *Chem. Phys.* **320**, 95 (2006).
  - [28] C. P. Koch, J. P. Palao, R. Kosloff, and F. Masnou-Seeuws, *Phys. Rev. A* **70**, 013402 (2004).
  - [29] S. J. Park, S. W. Suh, Y. S. Lee, and G. H. Jeung, *J. Molec. Spec.* **207**, 129 (2001).
  - [30] Y. S. Weinstein, M. A. Pravia, E. M. Fortunato, S. Lloyd, and D. G. Cory, *Phys. Rev. Lett.* **86**, 1889 (2001).
  - [31] B. Misra and E. C. G. Sudarshan, *J. Math. Phys.* **18**, 756 (1977).